



## SEQUENCE LISTING

<110> Kosan Biosciences, Inc.  
Santi, Daniel  
Peck, Larry  
Dayem, Linda  
Kealey, James

<120> HETEROLOGOUS PRODUCTION OF POLYKETIDES

<130> 30062-20049.00

<140> US 09/699,136

<141> 2000-10-27

<150> US 60/161,414

<151> 1999-10-25

<150> US 60/161,703

<151> 1999-10-27

<150> US 60/206,082

<151> 2000-05-18

<160> 2

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 447

<212> DNA

<213> Artificial Sequence

<220>

<223> Isolated and recombinant form of the full  
epimerase gene sequence

<221> CDS

<222> (1)...(444)

<400> 1

atg agt aat gag gat ctt ttc atc tgt atc gat cac gtg gca tat gcg	48
Met Ser Asn Glu Asp Leu Phe Ile Cys Ile Asp His Val Ala Tyr Ala	
1 5 10 15	

tgc ccc gac gcc gac gag gct tcc aag tac tac cag gag acc ttc ggc	96
Cys Pro Asp Ala Asp Glu Ala Ser Lys Tyr Tyr Gln Glu Thr Phe Gly	
20 25 30	

tgg cat gag ctc cac cgc gag gag aac ccg gag cag gga gtc gtc gag	144
Trp His Glu Leu His Arg Glu Glu Asn Pro Glu Gln Gly Val Val Glu	
35 40 45	

atc atg atg gcc ccg gct gcg aag ctg acc gag cac atg acc cag gtt	192
Ile Met Met Ala Pro Ala Ala Lys Leu Thr Glu His Met Thr Gln Val	

50	55	60	
cag gtc atg gcc ccg ctc aac gac gag tcg acc gtt gcc aag tgg ctt			240
Gln Val Met Ala Pro Leu Asn Asp Glu Ser Thr Val Ala Lys Trp Leu			
65	70	75	80
gcc aag cac aat ggt cgc gcc gga ctg cac cac atg gca tgg cgt gtc			288
Ala Lys His Asn Gly Arg Ala Gly Leu His His Met Ala Trp Arg Val			
85	90	95	
gat gac atc gac gcc gtc agc gcc acc ctg cgc gag cgc gcc gtg cag			336
Asp Asp Ile Asp Ala Val Ser Ala Thr Leu Arg Glu Arg Gly Val Gln			
100	105	110	
ctg ctg tat gac gag ccc aag ctc ggc acc ggc gcc aac cgc atc aac			384
Leu Leu Tyr Asp Glu Pro Lys Leu Gly Thr Gly Gly Asn Arg Ile Asn			
115	120	125	
ttc atg cat ccc aag tcg ggc aag ggc gtg ctc atc gag ctc acc cag			432
Phe Met His Pro Lys Ser Gly Lys Gly Val Leu Ile Glu Leu Thr Gln			
130	135	140	
tac ccg aag aac tga			447
Tyr Pro Lys Asn			
145			

<210> 2  
 <211> 148  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Deduced amino acid sequence of the epimerase gene  
 sequence

<400> 2															
Met	Ser	Asn	Glu	Asp	Leu	Phe	Ile	Cys	Ile	Asp	His	Val	Ala	Tyr	Ala
1			5					10					15		
Cys	Pro	Asp	Ala	Asp	Glu	Ala	Ser	Lys	Tyr	Tyr	Gln	Glu	Thr	Phe	Gly
	20						25				30				
Trp	His	Glu	Leu	His	Arg	Glu	Glu	Asn	Pro	Glu	Gln	Gly	Val	Val	Glu
	35					40				45					
Ile	Met	Met	Ala	Pro	Ala	Ala	Lys	Leu	Thr	Glu	His	Met	Thr	Gln	Val
	50				55				60						
Gln	Val	Met	Ala	Pro	Leu	Asn	Asp	Glu	Ser	Thr	Val	Ala	Lys	Trp	Leu
65				70				75					80		
Ala	Lys	His	Asn	Gly	Arg	Ala	Gly	Leu	His	His	Met	Ala	Trp	Arg	Val
			85				90				95				
Asp	Asp	Ile	Asp	Ala	Val	Ser	Ala	Thr	Leu	Arg	Glu	Arg	Gly	Val	Gln
	100						105				110				
Leu	Leu	Tyr	Asp	Glu	Pro	Lys	Leu	Gly	Thr	Gly	Gly	Asn	Arg	Ile	Asn
	115				120						125				
Phe	Met	His	Pro	Lys	Ser	Gly	Lys	Gly	Val	Leu	Ile	Glu	Leu	Thr	Gln
130					135					140					
Tyr	Pro	Lys	Asn												
145															